Derivation of Prediction Formulae

S. R. Searle

Biometrics Unit, Cornell University, Ithaca, N.Y.

BU-482-M August 1973

Abstract

This paper gives detailed derivation of formulae presented in "Prediction, mixed models and variance components", Searle [1974], and acts as an appendix thereto.

1A Introduction

Although many of the formulae in Searle [1974] are well-known results and their derivation is quite straightforward, those derivations are brought together here for the sake of completeness. Notation, paragraphing and equation numbers are in complete agreement with Searle [1974].

2.1A Best prediction

Derivation

Minimize, for $\underset{\sim}{A}$ positive definite and symmetric,

$$E(\tilde{\underset{\sim}{u}} - \underset{\sim}{u})'\underset{\sim}{A}(\tilde{\underset{\sim}{u}} - \underset{\sim}{u}) = \iint (\tilde{\underset{\sim}{u}} - \underset{\sim}{u})'\underset{\sim}{A}(\tilde{\underset{\sim}{u}} - \underset{\sim}{u}) f(\underset{\sim}{u},\underset{\sim}{y})\, d\underset{\sim}{u}\, d\underset{\sim}{y}$$

$$= \int \left[\int (\tilde{\underset{\sim}{u}} - \underset{\sim}{u})'\underset{\sim}{A}(\tilde{\underset{\sim}{u}} - \underset{\sim}{u}) f(\underset{\sim}{u}|\underset{\sim}{y})\, d\underset{\sim}{u} \right] f(\underset{\sim}{y})\, d\underset{\sim}{y}$$

where $f(\underset{\sim}{u}|\underset{\sim}{y})$ and $f(\underset{\sim}{y})$ are conditional and marginal densities respectively. Minimizing with respect to $\tilde{\underset{\sim}{u}}$ only requires minimizing of the integral over $\underset{\sim}{u}$ and gives

$$\int (2\underset{\sim}{A}\tilde{\underset{\sim}{u}} - 2\underset{\sim}{A}\underset{\sim}{u}) f(\underset{\sim}{u}|\underset{\sim}{y})\, d\underset{\sim}{u} = 0 \ .$$

Hence, since $\underset{\sim}{A}$ is positive definite,

$$\tilde{\underset{\sim}{u}} = \frac{\int \underset{\sim}{u} f(\underset{\sim}{u}|\underset{\sim}{y})\, d\underset{\sim}{u}}{\int f(\underset{\sim}{u}|\underset{\sim}{y})\, d\underset{\sim}{u}} = \int \underset{\sim}{u} f(\underset{\sim}{u}|\underset{\sim}{y})\, d\underset{\sim}{u} = E(\underset{\sim}{u}|\underset{\sim}{y}) \ . \qquad (3)$$

Expectation

$$E(\tilde{\underset{\sim}{u}}) = E_Y E_{U|Y}\,[E(\underset{\sim}{u}|\underset{\sim}{y})] = E_Y[E(\underset{\sim}{u}|\underset{\sim}{y})]$$

$$= \int \left[\int \underset{\sim}{u} f(\underset{\sim}{u}|\underset{\sim}{y})\, d\underset{\sim}{u} \right] f(\underset{\sim}{y})\, d\underset{\sim}{y}$$

$$= \iint \underset{\sim}{u} f(\underset{\sim}{u},\underset{\sim}{y})\, d\underset{\sim}{u}\, d\underset{\sim}{y} = E(\underset{\sim}{u}) \ . \qquad (4)$$

Variances and covariances

$$\mathrm{var}(\tilde{\underline{u}} - \underline{u}) = E(\tilde{\underline{u}} - \underline{u})(\tilde{\underline{u}} - \underline{u})', \text{ because } E(\tilde{\underline{u}} - \underline{u}) = \underline{0} \text{ from (4),}$$

$$= E_Y E_{U|Y}[E(\underline{u}|\underline{y})E(\underline{u}|\underline{y})' + \underline{u}\underline{u}' - \underline{u}E(\underline{u}|\underline{y})' - E(\underline{u}|\underline{y})\underline{u}']$$

$$= E_Y[E(\underline{u}|\underline{y})E(\underline{u}|\underline{y})' + E(\underline{u}\underline{u}'|\underline{y}) - 2E(\underline{u}|\underline{y})E(\underline{u}|\underline{y})']$$

$$= E_Y[E(\underline{u}\underline{u}'|\underline{y}) - E(\underline{u}|\underline{y})E(\underline{u}|\underline{y})']$$

$$= E_Y[\mathrm{var}(\underline{u}|\underline{y})] \,. \tag{5}$$

$$\mathrm{cov}(\tilde{\underline{u}}, \underline{u}') = E(\tilde{\underline{u}}\underline{u}') - E(\tilde{\underline{u}})E(\underline{u}')$$

$$= E_Y E_{U|Y}[E(\underline{u}|\underline{y})\underline{u}'] - E(\underline{u})E(\underline{u}')$$

$$= E_Y[E(\underline{u}|\underline{y})E(\underline{u}|\underline{y})'] - [E_Y E(\underline{u}|\underline{y})][E_Y E(\underline{u}|\underline{y})]'$$

$$= \mathrm{var}[E(\underline{u}|\underline{y})] = \mathrm{var}(\tilde{\underline{u}}) \tag{6}$$

$$\mathrm{cov}(\tilde{\underline{u}}, \underline{y}') = E(\tilde{\underline{u}}\underline{y}') - E(\tilde{\underline{u}})E(\underline{y}')$$

$$= E_Y E_{U|Y}[E(\underline{u}|\underline{y})\underline{y}'] - E(\underline{u})E(\underline{y}')$$

$$= E_Y E_{U|Y}(\underline{u}\underline{y}') - E(\underline{u})E(\underline{y}')$$

$$= E(\underline{u}\underline{y}') - E(\underline{u})E(\underline{y}')$$

$$= \mathrm{cov}(\underline{u}, \underline{y}') \tag{7}$$

Maximum correlation

As a function of $\underline{y}$ let p be any predictor of u, an element of $\underline{u}$. Then

$$\mathrm{cov}(p, u) = E\{[p - E(p)][u - E(u)]\}$$

$$= E\{[p - E(p)]u\}$$

$$= E_Y E_{U|Y}\{[p - E(p)]u\}$$

$$= E_Y\{[p - E(p)]E(u|\underset{\sim}{y})\} \text{ because } p \text{ is a function of } \underset{\sim}{y}$$

$$= E_Y\{[p - E(p)]\tilde{u}\}$$

$$= \text{cov}(p, \tilde{u}).$$

When $p = \tilde{u}$, $\text{cov}(\tilde{u}, u) = \text{cov}(\tilde{u}, \tilde{u}) = \sigma^2_{\tilde{u}}$

and

$$\rho(\tilde{u}, u) = \frac{\text{cov}(\tilde{u}, u)}{\sigma_{\tilde{u}}\sigma_u} = \frac{\sigma_{\tilde{u}}}{\sigma_u} \quad . \tag{8}$$

Hence in general

$$\rho^2(p, u) = \frac{\text{cov}^2(p, u)}{\sigma^2_p \sigma^2_u} = \frac{\text{cov}^2(p, \tilde{u})}{\sigma^2_p \sigma^2_{\tilde{u}}} \frac{\sigma^2_{\tilde{u}}}{\sigma^2_u} = \rho^2(p, \tilde{u})\rho^2(\tilde{u}, u) \ .$$

For choice of p this is maximum when $\rho^2(p, \tilde{u}) = 1$, i.e., $p = \tilde{u}$. Hence (8) is maximum $\rho(\tilde{u}, u)$. This proof follows Rao [1965, p. 221].

2.2A Best linear prediction

Matrix results

When $\text{tr}(\underset{\sim}{X}\underset{\sim}{P})$ exists its value is $\Sigma\Sigma x_{ij}p_{ji}$. Hence $\frac{\partial}{\partial x_{ij}}\text{tr}(\underset{\sim}{X}\underset{\sim}{P}) = p_{ji}$

and so
$$\frac{\partial}{\partial \underset{\sim}{X}}\text{tr}(\underset{\sim}{X}\underset{\sim}{P}) \overset{\text{def}}{\equiv} \left\{ \frac{\partial}{\partial x_{ij}}\text{tr}(\underset{\sim}{X}\underset{\sim}{P}) \right\} = \left\{p_{ji}\right\} = \underset{\sim}{P}' \ .$$

Also, because $\text{tr}(\underset{\sim}{X}\underset{\sim}{P}) = \text{tr}(\underset{\sim}{P}\underset{\sim}{X})$

$$\frac{\partial}{\partial \underset{\sim}{X}}\text{tr}(\underset{\sim}{X}\underset{\sim}{P}) = \frac{\partial}{\partial \underset{\sim}{X}}\text{tr}(\underset{\sim}{P}\underset{\sim}{X}) = \underset{\sim}{P}' \ . \tag{A1}$$

And since $\text{tr}(\underset{\sim}{X}'\underset{\sim}{P}) = \text{tr}(\underset{\sim}{P}'\underset{\sim}{X})$

$$\frac{\partial}{\partial \underset{\sim}{X}}\text{tr}(\underset{\sim}{X}'\underset{\sim}{P}) = \frac{\partial}{\partial \underset{\sim}{X}}\text{tr}(\underset{\sim}{P}'\underset{\sim}{X}) = \underset{\sim}{P} \ . \tag{A2}$$

Hence

$$\frac{\partial}{\partial \underset{\sim}{X}} tr(\underset{\sim}{P}_1 \underset{\sim}{X}' \underset{\sim}{Q}\underset{\sim}{X}\underset{\sim}{P}_2) = \underset{\sim}{Q}\underset{\sim}{X}\underset{\sim}{P}_2\underset{\sim}{P}_1 + \underset{\sim}{Q}'\underset{\sim}{X}\underset{\sim}{P}_1'\underset{\sim}{P}_2' \quad . \tag{A3}$$

Also

$$\frac{\partial}{\partial \underset{\sim}{X}} tr(\underset{\sim}{P}_1\underset{\sim}{P}_2\underset{\sim}{X}\underset{\sim}{Q}_1\underset{\sim}{Q}_2) = \frac{\partial}{\partial \underset{\sim}{X}} tr(\underset{\sim}{X}\underset{\sim}{Q}_1\underset{\sim}{Q}_2\underset{\sim}{P}_1\underset{\sim}{P}_2) = \underset{\sim}{P}_2'\underset{\sim}{P}_1'\underset{\sim}{Q}_2'\underset{\sim}{Q}_1' \tag{A4}$$

and

$$\frac{\partial}{\partial \underset{\sim}{X}}(\underset{\sim}{r}'\underset{\sim}{P}\underset{\sim}{X}\underset{\sim}{Q}\underset{\sim}{s}) = \frac{\partial}{\partial \underset{\sim}{X}} tr(\underset{\sim}{r}'\underset{\sim}{P}\underset{\sim}{X}\underset{\sim}{Q}\underset{\sim}{s}) = \underset{\sim}{P}'\underset{\sim}{r}\underset{\sim}{s}'\underset{\sim}{Q} \; . \tag{A5}$$

<u>Minimization</u>

For $\tilde{\underset{\sim}{u}} = \underset{\sim}{a} + \underset{\sim}{B}\underset{\sim}{y}$ we minimize, for positive definite symmetric $\underset{\sim}{A}$,

$$E(\tilde{\underset{\sim}{u}} - \underset{\sim}{u})'\underset{\sim}{A}(\tilde{\underset{\sim}{u}} - \underset{\sim}{u}) = E(\underset{\sim}{a} + \underset{\sim}{B}\underset{\sim}{y} - \underset{\sim}{u})'\underset{\sim}{A}(\underset{\sim}{a} + \underset{\sim}{B}\underset{\sim}{y} - \underset{\sim}{u})$$

$$= E\left\{\underset{\sim}{a}'\underset{\sim}{A}\underset{\sim}{a} + 2\underset{\sim}{a}'\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{y} - 2\underset{\sim}{a}'\underset{\sim}{A}\underset{\sim}{u} + (\underset{\sim}{u}' \;\; \underset{\sim}{y}')\begin{bmatrix} -\underset{\sim}{I} \\ \underset{\sim}{B}' \end{bmatrix}\underset{\sim}{A}(-\underset{\sim}{I} \;\; \underset{\sim}{B})\begin{bmatrix} \underset{\sim}{u} \\ \underset{\sim}{y} \end{bmatrix}\right\}$$

$$= \underset{\sim}{a}'\underset{\sim}{A}\underset{\sim}{a} + 2\underset{\sim}{a}'\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{\mu}_Y - 2\underset{\sim}{a}'\underset{\sim}{A}\underset{\sim}{\mu}_U$$

$$+ (\underset{\sim}{\mu}_U' \;\; \underset{\sim}{\mu}_Y')\begin{bmatrix} -\underset{\sim}{A} & -\underset{\sim}{A}\underset{\sim}{B} \\ -\underset{\sim}{B}'\underset{\sim}{A} & \underset{\sim}{B}'\underset{\sim}{A}\underset{\sim}{B} \end{bmatrix}\begin{bmatrix} \underset{\sim}{\mu}_U \\ \underset{\sim}{\mu}_Y \end{bmatrix} + tr\begin{bmatrix} -\underset{\sim}{A} & -\underset{\sim}{A}\underset{\sim}{B} \\ -\underset{\sim}{B}'\underset{\sim}{A} & \underset{\sim}{B}'\underset{\sim}{A}\underset{\sim}{B} \end{bmatrix}\begin{bmatrix} \underset{\sim}{V}_U & \underset{\sim}{C} \\ \underset{\sim}{C}' & \underset{\sim}{V} \end{bmatrix}$$

$$= \underset{\sim}{a}'\underset{\sim}{A}\underset{\sim}{a} + 2\underset{\sim}{a}'\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{\mu}_Y - 2\underset{\sim}{a}'\underset{\sim}{A}\underset{\sim}{\mu}_U$$

$$- \underset{\sim}{\mu}_U'\underset{\sim}{A}\underset{\sim}{\mu}_U - 2\underset{\sim}{\mu}_U'\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{\mu}_Y + \underset{\sim}{\mu}_Y'\underset{\sim}{B}'\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{\mu}_Y$$

$$+ tr(-\underset{\sim}{A}\underset{\sim}{V}_U - \underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{C}' - \underset{\sim}{B}'\underset{\sim}{A}\underset{\sim}{C} + \underset{\sim}{B}'\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{V}) \; . \tag{A6}$$

Using (A1) - (A5) to differentiate this with respect to elements of $\underline{B}$ gives

$$2\underline{A}'\underline{a}\underline{\mu}_Y' - 2\underline{A}'\underline{\mu}_U\underline{\mu}_Y' + \underline{A}\underline{B}\underline{\mu}_Y\underline{\mu}_Y' + \underline{A}'\underline{B}\underline{\mu}_Y'\underline{\mu}_Y - \underline{A}'\underline{C} - \underline{A}\underline{C} + \underline{A}\underline{B}\underline{V} + \underline{A}'\underline{B}\underline{V}' = \underline{0} .$$

The symmetry of $\underline{A}$ and $\underline{V}$ and the non-singularity of $\underline{A}$ reduce this to

$$(\underline{a} - \underline{\mu}_U + \underline{B}\underline{\mu}_Y)\underline{\mu}_Y' + \underline{B}\underline{V} = \underline{C} . \tag{A7}$$

Differentiating (A6) with respect to $\underline{a}$ gives $\underline{a} = \underline{\mu}_U - \underline{B}\underline{\mu}_Y$ so that (A7) gives $\underline{B} = \underline{C}\underline{V}^{-1}$ and hence

$$\tilde{\underline{u}} = \underline{a} + \underline{B}\underline{y} = \underline{\mu}_U + \underline{C}\underline{V}^{-1}(\underline{y} - \underline{\mu}_Y) \tag{16}$$

2.4A Mixed model prediction

With $\underline{w} = \underline{K}'\underline{\beta} + \underline{u}$ and $\tilde{\underline{w}} = \underline{B}\underline{y}$ for $\underline{B}\underline{X} = \underline{K}'$ we have

$$E\begin{bmatrix} \underline{w} \\ \underline{y} \end{bmatrix} = \begin{bmatrix} \underline{K}'\underline{\beta} \\ \underline{X}\underline{\beta} \end{bmatrix} \quad \text{and} \quad \text{var}\begin{bmatrix} \underline{w} \\ \underline{y} \end{bmatrix} = \begin{bmatrix} \underline{V}_U & \underline{C} \\ \underline{C}' & \underline{V} \end{bmatrix} ,$$

and seek to minimize $E(\tilde{\underline{w}} - \underline{w})'\underline{A}(\tilde{\underline{w}} - \underline{w})$ for $\underline{A}$ being positive definite symmetric. We minimize

$$\lambda = E(\tilde{\underline{w}} - \underline{w})'\underline{A}(\tilde{\underline{w}} - \underline{w}) + \text{tr}[\underline{T}(\underline{B}\underline{X} - \underline{K}')]$$

where $\underline{T}$ is a matrix of Lagrange multipliers. Since $\underline{A}$ is positive definite there is no loss of generality in writing $\underline{T} = 2\underline{M}\underline{A}$ for some matrix $\underline{M}$. Then

$$\lambda = E(\underline{B}\underline{y} - \underline{w})'\underline{A}(\underline{B}\underline{y} - \underline{w}) + 2\text{tr}[\underline{M}\underline{A}(\underline{B}\underline{X} - \underline{K}')]$$

$$= E(\underset{\sim}{w}'\ \underset{\sim}{y}')\begin{bmatrix} -\underset{\sim}{I} \\ \underset{\sim}{B} \end{bmatrix}\underset{\sim}{A}(-\underset{\sim}{I}\ \ \underset{\sim}{B})\begin{bmatrix} \underset{\sim}{w} \\ \underset{\sim}{y} \end{bmatrix} + 2\mathrm{tr}[\underset{\sim}{MA}(\underset{\sim}{BX} - \underset{\sim}{K}')]$$

$$= (\underset{\sim}{\beta}'\underset{\sim}{K}\ \ \underset{\sim}{\beta}'\underset{\sim}{X}')\begin{bmatrix} -\underset{\sim}{A} & -\underset{\sim}{AB} \\ -\underset{\sim}{B}'\underset{\sim}{A} & \underset{\sim}{B}'\underset{\sim}{AB} \end{bmatrix}\begin{bmatrix} \underset{\sim}{K}'\underset{\sim}{\beta} \\ \underset{\sim}{X\beta} \end{bmatrix} + \mathrm{tr}\begin{bmatrix} -\underset{\sim}{A} & -\underset{\sim}{AB} \\ -\underset{\sim}{B}'\underset{\sim}{A} & \underset{\sim}{B}'\underset{\sim}{AB} \end{bmatrix}\begin{bmatrix} \underset{\sim}{V}_U & \underset{\sim}{C} \\ \underset{\sim}{C}' & \underset{\sim}{V} \end{bmatrix} + 2\mathrm{tr}[\underset{\sim}{MA}(\underset{\sim}{BX} - \underset{\sim}{K}')]$$

$$= -\underset{\sim}{\beta}'\underset{\sim}{KAK}'\underset{\sim}{\beta} - 2\underset{\sim}{\beta}'\underset{\sim}{KABX\beta} + \underset{\sim}{\beta}'\underset{\sim}{X}'\underset{\sim}{B}'\underset{\sim}{ABX\beta}$$

$$+ \mathrm{tr}(-\underset{\sim}{AV}_U - \underset{\sim}{ABC}' + \underset{\sim}{B}'\underset{\sim}{AC} + \underset{\sim}{B}'\underset{\sim}{ABV}) + 2\mathrm{tr}[\underset{\sim}{MA}(\underset{\sim}{BX} - \underset{\sim}{K}')].$$

Using the matrix results (A1) through (A5) gives $\partial\lambda/\partial\underset{\sim}{B} = \underset{\sim}{0}$ as (using $\underset{\sim}{A} = \underset{\sim}{A}'$)

$$-2\underset{\sim}{AK}'\underset{\sim}{\beta\beta}'\underset{\sim}{X}' + \underset{\sim}{ABX\beta\beta}'\underset{\sim}{X}' + \underset{\sim}{ABX\beta\beta}'\underset{\sim}{X}' - \underset{\sim}{AC} - \underset{\sim}{AC} + \underset{\sim}{ABV} + \underset{\sim}{ABV} + 2\underset{\sim}{AM}'\underset{\sim}{X}' = \underset{\sim}{0}. \quad \text{(A8)}$$

And $\partial\lambda/\partial\underset{\sim}{M} = \underset{\sim}{0}$ gives $\underset{\sim}{BX} = \underset{\sim}{K}'$ because $\underset{\sim}{A}$ is non-singular. Using these results in (A7) gives

$$-\underset{\sim}{BX\beta\beta}'\underset{\sim}{X}' + \underset{\sim}{BX\beta\beta}'\underset{\sim}{X}' - \underset{\sim}{C} + \underset{\sim}{BV} + \underset{\sim}{M}'\underset{\sim}{X}' = \underset{\sim}{0},$$

i.e.

$$\underset{\sim}{BV} + \underset{\sim}{M}'\underset{\sim}{X}' = \underset{\sim}{C}. \qquad \text{(A9)}$$

This and

$$\underset{\sim}{BX} = \underset{\sim}{K}' \qquad \text{(A10)}$$

are the equations to be solved for $\underset{\sim}{B}$. From (A9)

$$\underset{\sim}{B} = (\underset{\sim}{C} - \underset{\sim}{M}'\underset{\sim}{X}')\underset{\sim}{V}^{-1} \qquad \text{(A11)}$$

and substitution in (A10) gives

$$\underset{\sim}{M}' = (\underset{\sim}{CV}^{-1}\underset{\sim}{X} - \underset{\sim}{K}')(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^-$$

so that in (A11)

$$\underset{\sim}{B} = \underset{\sim}{CV}^{-1} + (\underset{\sim}{K}' - \underset{\sim}{CV}^{-1}\underset{\sim}{X})(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^-\underset{\sim}{X}'\underset{\sim}{V}^{-1}\ . \qquad (31)$$

Variances and covariances

$$X\beta^{o} = X(X'V^{-1}X)^{-}X'V^{-1}y$$

$$\begin{aligned}
\mathrm{var}(X\beta^{o}) &= X(X'V^{-1}X)^{-}X'V^{-1}X(X'V^{-1}X)^{-\prime}X' \\
&= X(X'V^{-1}X)^{-}X'V'X(X'V^{-1}X)^{-\prime}X'V^{-\frac{1}{2}}V^{\frac{1}{2}} \text{ because } V \text{ is positive definite,} \\
&= X(X'V^{-1}X)^{-}(X'V^{-\frac{1}{2}})V^{\frac{1}{2}}, \text{ because } Q'Q(Q'Q)^{-}Q' = Q' \text{ for any } Q, \\
&= X(X'V^{-1}X)^{-}X' \;.
\end{aligned}$$

$$K' = BX$$

$$\begin{aligned}
\mathrm{var}(K'\beta^{o}) &= BX(X'V^{-1}X)^{-}X'B' \\
&= K'(X'V^{-1}X)K \qquad (35)
\end{aligned}$$

$$\begin{aligned}
\mathrm{cov}(X\beta^{o}, y') &= X(X'V^{-1}X)^{-}X'V^{-1}V \\
&= X(X'V^{-1}X)^{-}X' \\
&= \mathrm{var}(X\beta^{o})
\end{aligned}$$

$$\tilde{u}^{o} = CV^{-1}(y - X\beta^{o})$$

$$\begin{aligned}
\mathrm{var}(\tilde{u}^{o}) &= CV^{-1}VV^{-1}C' + CV^{-1}[-\,\mathrm{var}(X\beta^{o})]V^{-1}C' \\
&= CV^{-1}C' - CV^{-1}X(X'V^{-1}X)^{-}X'V^{-1}C \qquad (36)
\end{aligned}$$

$$\begin{aligned}
\mathrm{cov}(X\beta^{o}, \tilde{u}^{o\prime}) &= \mathrm{cov}[X\beta^{o},\ (y - X\beta^{o})'V^{-1}C'] \\
&= [\mathrm{var}(X\beta^{o}) - \mathrm{var}(X\beta^{o})]V^{-1}C \\
&= 0 \qquad (37)
\end{aligned}$$

$$\tilde{w} = K\beta^{o} + u^{o}$$

$$\mathrm{var}(\tilde{w}) = \mathrm{var}(K\beta^{o}) + \mathrm{var}(u^{o}) \tag{38}$$

$$\mathrm{cov}(\tilde{u}^{o}, u') = \mathrm{cov}\{CV^{-1}[I - X(X'V^{-1}X)^{-}X'V^{-1}]y, u'\}$$

$$= CV^{-1}[I - X(X'V^{-1}X)^{-}X'V^{-1}]C'$$

$$= \mathrm{var}(u^{o}) \tag{39}$$

$$\mathrm{var}(\tilde{u}^{o} - u) = \mathrm{var}(u) - \mathrm{var}(u^{o})$$

$$= V_{U} - \mathrm{var}(u^{o}) \tag{40}$$

$$\mathrm{cov}(K'\beta^{o}, u') = \mathrm{cov}(BX\beta^{o}, u')$$

$$= BX(X'V^{-1}X)^{-}X'V^{-1}C' \tag{41}$$

$$\mathrm{var}(\tilde{w} - w) = \mathrm{var}(K'\beta^{o} + \tilde{u}^{o} - K'\beta - u)$$

$$= \mathrm{var}(K'\beta^{o}) + \mathrm{var}(u^{o} - u) + \mathrm{cov}[K'\beta^{o}, (u^{o} - u)'] + \mathrm{cov}[(u^{o} - u), (K'\beta^{o})']$$

$$= \mathrm{var}(K'\beta^{o}) + \mathrm{var}(u^{o} - u) - \mathrm{cov}(K'\beta^{o}, u') - \mathrm{cov}\,(u, \beta^{o\prime}K) \tag{42}$$

3.1A Calculating the predictor (in the mixed model)

From the second equation of

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}D \\ Z'R^{-1}X & Z'R^{-1}Z + D^{-1} \end{bmatrix} \begin{bmatrix} \beta^{o} \\ u^{o} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix} \tag{A12}$$

we get

$$u^{o} = (Z'R^{-1}Z + D^{-1})^{-1}Z'R^{-1}(y - X\beta^{o}) \tag{A13}$$

So long as $\mathbf{D} = \text{var}(\mathbf{u})$ is non-singular $(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})^{-1}$ always exists, because $\mathbf{R}^{-1}$ and $\mathbf{D}^{-1}$ are symmetric, and equal to $\mathbf{P}'\mathbf{P}$ and $\mathbf{Q}^{-1'}\mathbf{Q}^{-1}$ say, and so

$$\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1} = \mathbf{Q}^{-1'}(\mathbf{K}'\mathbf{K} + \mathbf{I})\mathbf{Q}^{-1} \text{ for } \mathbf{K} = \mathbf{PZQ}$$

is non-singular (Searle [1971, p. 24, lemma 8]). (A13) always holds, therefore, and substituting it into the first equation of (A12) gives, after a little reduction,

$$\mathbf{X}'\mathbf{W}\mathbf{X}\beta^{o} = \mathbf{X}'\mathbf{W}\mathbf{y} \tag{A14}$$

where

$$\mathbf{W} = \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1} .$$

It remains to show that $\mathbf{WV} = \mathbf{I}$ which it does:

$$\begin{aligned}
\mathbf{WV} &= [\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}](\mathbf{ZDZ}' + \mathbf{R}) \\
&= \mathbf{R}^{-1}\mathbf{ZDZ}' + \mathbf{I} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})^{-1}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{ZDZ}' + \mathbf{Z}') \\
&= \mathbf{R}^{-1}\mathbf{ZDZ}' + \mathbf{I} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})^{-1}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})\mathbf{DZ}' \\
&= \mathbf{R}^{-1}\mathbf{ZDZ}' + \mathbf{I} - \mathbf{R}^{-1}\mathbf{ZDZ}' \\
&= \mathbf{I} .
\end{aligned}$$

With $\mathbf{W}$ and $\mathbf{V}$ both being symmetric this implies $\mathbf{W} = \mathbf{V}^{-1}$. Hence (A14) is $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\beta^{o} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ for which $\beta^{o} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ is a familiar solution, as in (50). It remains to show that $\mathbf{u}^{o}$ of (A13) is $\mathbf{u}^{o} = \mathbf{DZ}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\beta^{o})$ of (51). It is, because in (A13)

$$\begin{aligned}
(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1} &= (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{VV}^{-1} \\
&= (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}(\mathbf{ZDZ}' + \mathbf{R})\mathbf{V}^{-1} \\
&= (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})^{-1}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})\mathbf{DZ}'\mathbf{V}^{-1} \\
&= \mathbf{DZ}'\mathbf{V}^{-1} .
\end{aligned}$$

ACKNOWLEDGEMENT

Preparation of the paper was partially supported by Grant GJ 31746 from the National Science Foundation, Washington, D.C.

REFERENCES

Rao, C. R. [1965]. Linear Statistical Inference and its Applications. Wiley, New York.

Searle, S. R. [1971]. Linear Models. Wiley, New York.

Searle, S. R. [1974]. "Prediction, mixed models and variance components", in Proceedings of a Conference on Reliability and Biometry, Tallahassee, Florida, 1973, Eds. F. Proschan and R. J. Serfling, a S.I.A.M. publication.